The EMBO Journal vol.10 no.7 pp.1809 - 1816, 1991

A T-cell specific TCR δ DNA binding protein is a member of the human GATA family

Virginie Joulin, Dominique Bories, Jean-François Eléouet, Marie-Claude Labastie, Stany Chrétien[1], Marie-Geneviève Mattéi[2] and Paul-Henri Roméo

INSERM U.91, Hôpital Henri Mondor, 94010 Créteil, [1]INTS, 6 rue Alexandre Cabanel, 75015 Paris and [3]INSERM U. 242, Hôpital de la Timone, 13385 Marseille, France

Communicated by M.Yaniv

The human GATA1, hGATA1 (previously called NF-E1, GF-1 or Eryf-1), a major sequence-specific DNA-binding protein of the erythrocytic lineage, is a member of a zinc-finger family of DNA-binding proteins. We report here the cloning of a human cDNA for a new member of this family. This member, called hGATA3, has 85% amino acid homology with hGATA1 in the DNA-binding domain and no homology elsewhere in the protein. Unlike hGATA1, hGATA3 is not localized on the X chromosome and we map it to the 10p15 band of the human genome. Northern blot analysis indicates that this factor is a T-cell specific transcription factor, present before activation and up-regulated during T-cell activation. The encoded hGATA3 protein, made in an *in vitro* transcription-translation assay, binds the WGATAR motif present in the human T-cell receptor (TCR) δ gene enhancer and, by transfection in HeLa cells, we show that hGATA3 can transactivate this TCR δ gene enhancer. Interestingly this enhancer binds and is also transactivated by hGATA1. Conversely, the promoter of the human glycophorin B (GPB), which is erythroid-specific and contains two WGATAR motifs, binds and is transactivated by hGATA1 and, to a lesser extent, by hGATA3. These results indicate that the activation of specific genes by hGATA1 or hGATA3 is partly governed by the lineage expression of these two factors during haematopoiesis and that, in the T-cell lineage, hGATA3 binds the human TCR δ gene enhancer and is involved in its expression.

Key words: human GATA family/haematopoietic *trans*-acting factors/T-cell differentiation

Introduction

Terminal differentiation of the various haematopoietic lineages is characterized by the co-ordinated appearance of specific gene products. This phenomenon is partly regulated at the gene transcription level and therefore knowledge of the mechanisms that control the transcription of genes during terminal haematopoietic differentiation is essential. Although such mechanisms have begun to be defined at the molecular level for genes specifically transcribed in the erythroid lineage, relatively little is known about the mechanisms that regulate the expression of T-lymphocyte specific genes.

In the erythroid lineage, the major DNA-binding protein, hGATA1 [previously called NF-E1 (Wall *et al.*, 1988), GF-1 (Martin *et al.*, 1989) or Eryf-1 (Evans *et al.*, 1988)] has been cloned (Tsai *et al.*, 1989). This protein recognized a sequence WGATAR (W = T or A, R = G or A) found in the regulatory regions of most of the erythroid-specific genes. It is a member of a new zinc-finger family of DNA-binding proteins and its expression, previously thought to be restricted to erythroid cells (Tsai *et al.*, 1989), has recently been shown to occur also in megakaryocytic and mastocytic cells (Martin *et al.*, 1990; Romeo *et al.*, 1990). mGATA1 and cGATA1, the murine and chicken homologues of hGATA1, have been shown to be potent transactivators of artificial promoters containing multiple copies of a WGATAR sequence (Yamamoto *et al.*, 1990), and we have shown that mGATA1 was able to transactivate various erythroid-specific promoters (our unpublished data). All those data indicated that GATA1 might be a major transactivator of terminally expressed erythroid-specific genes.

Differentiation of the T-cell lymphoid lineage is also associated with the appearance of various specific markers. The majority of mature peripheral T-cells harbours the T-cell α/β complex receptor while a minority of T-cell expresses the TCR γ/δ complex. These TCR recognize foreign antigen in the context of self major histocompatibility complex (MHC) molecules. Analysis of the transcriptional regulatory regions of human TCR α, β and δ genes has shown that their expression was partly regulated by enhancers located 3′ to the initiation of the transcription of these genes (Ho *et al.*, 1989; Bories *et al.*, 1990; Gottschalk and Leiden, 1990; Redondo *et al.*, 1990). These enhancers conferred T-cell specificity to heterologous promoters and DNase I footprint analyses have shown the binding of numerous T-cell specific factors on them (Ho *et al.*, 1989; Gottschalk and Leiden, 1990; Redondo *et al.*, 1990). Interestingly, T-cell specific binding sites within the TCR δ (footprint δE4), TCR β (footprint Tβ2) and TCR α (footprint Tα3) enhancers contained sequences that perfectly matched the WGATAR consensus sequence. Furthermore, electrophoretic mobility shift assays have shown that the DNA region corresponding to footprint Tβ2 could bind a nuclear factor present either in T-cell lines or in an erythroleukaemic cell line but not in B or Hela cell lines (Gottschalk and Leiden, 1990).

In chicken, cGATA1 is a member of a multigene family where each member seems to display a unique pattern of tissue-specific expression. cGATA2 is expressed not only in erythroid cells but also in many non-erythroid tissues, whereas cGATA3 is expressed in adult erythrocytes, T-lymphocytes and embryonic brain (Yamamoto *et al.*, 1990). In human, a non-erythroid GATA-binding protein has recently been described as important for activity of the human preproendothelin-1 promoter in endothelial cells (Wilson *et al.*, 1990), indicating that a multigene GATA family is also present in human. In order to identify a member of this family that might interact with the WGATAR sequence present in the TCR α, β and/or δ enhancers and specific to T-lymphocytes, we have screened a human T-cell

cDNA library with a DNA probe corresponding to the DNA-binding domain of hGATA1 and cloned a new member of the human GATA family. Its structure, chromosomal localization, binding specificity and expression pattern were also studied. Finally, as hGATA3 and hGATA1 have no sequence homology outside the DNA-binding domain, we analysed their transactivating activities on erythroid and T-lymphoid gene regulatory regions.

Results

WGATAR motifs present in the human TCRδ gene enhancer confer T-cell enhanced transcription

Several recent reports have shown multiple T-cell specific footprints in WGATAR sequences present in human or mouse TCRδ (Ho *et al.*, 1989), β (Gottschalk *et al.*, 1990; Takeda *et al.*, 1990), γ (Spencer *et al.*, 1991) and δ (Redondo *et al.*, 1990) gene enhancers. To look for any function of this *cis*-acting sequence in T-cells, the δE4 region of the human TCRδ enhancer (Redondo *et al.*, 1990; Bories *et al.*, 1990) or the δE4 region mutated on the two WGATAR motifs were linked to a minimal promoter [−36 to +47] of the human glycophorin B (GPB) gene (Vignal *et al.*, 1990) upstream from the chloramphenicol acetyl transferase (CAT) reporter gene. This minimal GPB promoter did not contain any WGATAR motif and gave accurate initiation of transcription when linked to heterologous enhancer sequences (our unpublished data). These constructs were transfected into Jurkat cells and, as shown in Figure 1, the δE4 construct showed a five-fold enhanced transcriptional activity when compared with the minimal GPB promoter. Point mutations on the two WGATAR motifs completely abolished this enhanced transcription and, together with a WGATAR DNA-binding activity found in Jurkat cells (data not shown), these results indicated a WGATAR transcriptional activity in T-cells.

Cloning and structure of human GATA3

A PCR-amplified DNA fragment corresponding to the two zinc-finger domains of the human GATA1 cDNA was used to screen a human T-cell library and five overlapping clones were isolated. Based on the sequences of those five clones, we obtained the now named hGATA3 sequence [cGATA2 has already been described in the chicken (Yamatoto *et al.*, 1990)]. Analysis of this sequence (EMBL Data library accession number X58072), predicted a single open reading frame that started at position 152 and extended to position 1483. The 5′ non-coding region was 151 bp long whereas there were 869 bp of 3′ non-coding region [excluding the poly(A) tail]. The putative ATG translation start codon lay within the sequence 5′-GGCCATGG-3′ that perfectly matched the consensus start sequence 5′-C/G A/G CC ATGG-3′ proposed by Kozak (Kozak, 1984). It was preceded by several in-frame termination codons (our unpublished data) which argued strongly for this ATG to be the initiation methionine codon. The 3′ non-coding region was highly AT-rich (>60%) and contained several ATTTA sequences previously found in many unstable mRNAs (Kruys *et al.*, 1989). This feature is not found in hGATA1 mRNA and might be related to the cell-type distribution and function of hGATA3. When the amino acid sequence of hGATA3 was aligned with that of hGATA1 (Figure 2), a remarkable identity was found within the entire DNA-binding domain, with a higher homology in the carboxy-terminal finger (88%) compared with that in the amino-terminal finger (81%). Apart from this domain, no significant homology was found between the two proteins, suggesting that they might share only their DNA-binding properties.

Chromosomal localization of hGATA3

We determined the chromosomal localization of hGATA3 using *in situ* hybridization of an hGATA3 probe on spreads of metaphase chromosomes. The hGATA3 probe extended from nucleotide 200 to nucleotide 800 and thus did not contain the DNA-binding domain of hGATA3. In the 100 metaphase cells examined after *in situ* hybridization, there were 219 silver grains associated with chromosomes and 57 of these (26%) were located on chromosome 10 (Figure 3A). The distribution of grains on this chromosome was not random: 86% of them mapped to the (p14−p15) region of chromosome 10 with a maximum in the 10p15 band (Figure 3B). These results allow us to map the hGATA3 probe to the 10p15 band of the human genome.

Haematopoietic expression pattern of hGATA3

Since hGATA1 mRNA has been found in erythrocytic, megakaryocytic and mastocytic cells (Martin *et al.*, 1990; Romeo *et al.*, 1990), we have investigated the presence of hGATA3 mRNA in these lineages as well as in other

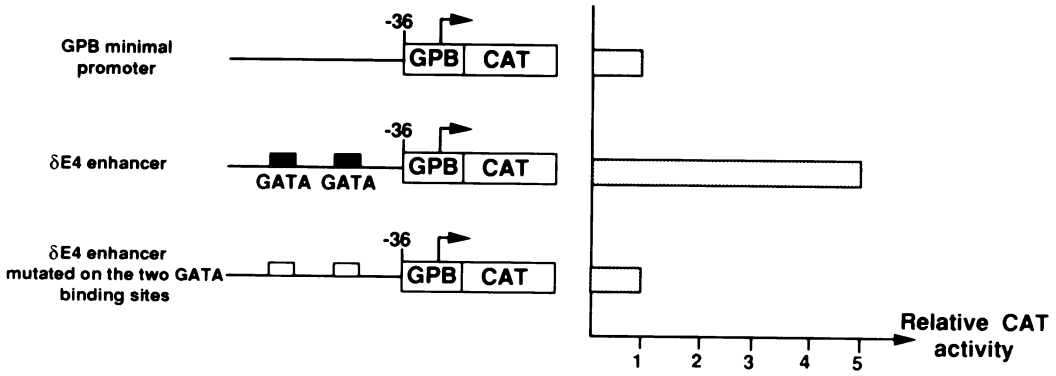

Fig. 1. The WGATAR motifs present in the δE4 region of the human TCRδ gene enhancer are necessary for enhanced transcriptional activity in T-cells. The three constructs shown were transfected in Jurkat cells together with a plasmid containing the RSV promoter driving the luciferase gene. After normalization of transfections, the transcriptional activity of each construct was monitored by CAT activities. The level of CAT expressed above the CAT reporter gene without promoter is indicated.

```
hGATA1  MEFPGLGSLGTSEPLP.................QFVDPALVSSTPESGVF 37
        **          *   *                  * *     *  *
hGATA3  MEVTADQPRWVSHHHPAVLNGQHPDTHHPGLSHSYMDAAQYPLPEEVDVL 50

hGATA1  FPSGPEGLDAAASSTAPSTATAAAAALAYYRDAEAYRHSPVFQVYPLLNC 83
        *    *              **          *      * *     *
hGATA3  FNIDGQGNHVPPYYGNSVRAT.......VQRYPPTHHGSQVCRPPLLHGS 93

hGATA1  MEGIPGG.........SPYAGWAYGKTGLYPASTVCPTREDSPPQAVEDL 124
          **          **       **     *          **     *
hGATA3  LPWLDGGKALGSHHTASPWNLSPFSKTSIHHGS..PGPLSVYPPASSSSL 141

hGATA1  DGKGSTSFLETL.....KTERLSPDLLTLGPA..............LP 153
         *     * *      *     * * * * *               **
hGATA3  SGGHASPHLFTFPPTPPKDVSPDPSLSTPGSAGSARQDEKECLKYQVPLP 191

hGATA1  SSLPVPNS.......AYGG..............PDFSSTFFSP....T 176
         *     *       * **                * **  * *
hGATA3  DSMKLESSHSRGSMTALGGASSSTHHPITTYPPYVPEYSSGLFPPSSLLG 241

hGATA1  GSPLNSAAYSSPKLRGTLPLPPCEARECVNCGATATPLWRRDRTGHYLCN 226
        ***       * ** *          ********** ******* *******
hGATA3  GSPTGFGCKSRPKAR......SSTGRECVNCGATSTPLWRRDGTGHYLCN 285

hGATA1  ACGLYHKMNGQNRPLIRPKKRLIVSKRAGTQCTNCQTTTTTLWRRNASGD 276
        ***************** ** **     **** * ************** **
hGATA3  ACGLYHKMNGQNRPLIKPKRRLSAARRAGTSCANCQTTTTTLWRRNANGD 335

hGATA1  PVCNACGLYYKLHQVNRPLTMRKDGIQTRNRKASGKGKKKRGSSLGGTGA 326
        *************  ****** * ******** * * **
hGATA3  PVCNACGLYYKLHNINRPLTMKKEGIQTRNRKMSSKSKK..........  374

hGATA1  AEGPAGGFMVVAGGSGSGNCGEVASGLTLGPPGTAHLYQGLGPVVLSGPV 376
                            * *   *   *         *  **
hGATA3  ..................CKKVHDSLEDFP.....KNSSFNPAALSRHM 400

hGATA1  SHLMPFPGPLLGSPTGSFPTGPMPPTTSTTVVAPLSS 413
        * *         *     **  **         **
hGATA3  SSLSHISPFSHSSHMLTTPTPMHPPSSLSFGPHHPSSMVTAMG 443
```

Fig. 2. Amino acid sequence comparison of hGATA1 and hGATA3. The amino acid sequence of hGATA3 was plotted against that of hGATA1. A single region of homology appeared which corresponded to the zinc finger region. Numbers on the right correspond to the position of the last amino acid of each line.

haematopoietic lineages. A single 3.1 kb mRNA was found by Northern analysis of RNAs from various haematopoietic lineages hybridized with a hGATA3 probe (Figure 4A). This mRNA was present in α/β or γ/δ T-cell lines (Jurkat or Peer) but not in cell lines derived from the other haematopoietic lineages (B cells, megakaryocytes, erythrocytes, histiocytes or promyelocytes). We next investigated the possible regulation of hGATA3 mRNA during the activation of Jurkat cell line. When Jurkat cells were treated with PMA plus PHA, a marked increase in the hGATA3 mRNA content was observed (Figure 4A, lanes 8 and 9). This mRNA increase was correlated with an increased WGATAR DNA binding activity of extracts from activated Jurkat cells (our unpublished data). Finally, we performed a Northern blot analysis of mRNA isolated either from resting T-lymphocytes or from PMA/PHA-activated T-lymphocytes. As for the Jurkat cell line, an increase in the amount of hGATA3 mRNA was found during T lymphocyte activation, suggesting a possible function of this factor during T-cell activation (Figure 4B).

hGATA3 binds to the WGATAR sequence present in the human TCRδ enhancer

hGATA3 protein has a zinc-finger domain highly homologous to hGATA1, and therefore we have investigated its DNA-binding activity. The human cDNA clone was transcribed *in vitro* and the RNA translated in a rabbit reticulocyte lysate in the presence of [^{35}S]methionine. The synthesized protein migrated with a relative molecular mass of 49 000 when analysed by SDS−PAGE (Figure 5A). This agreed with the expected molecular weight deduced from

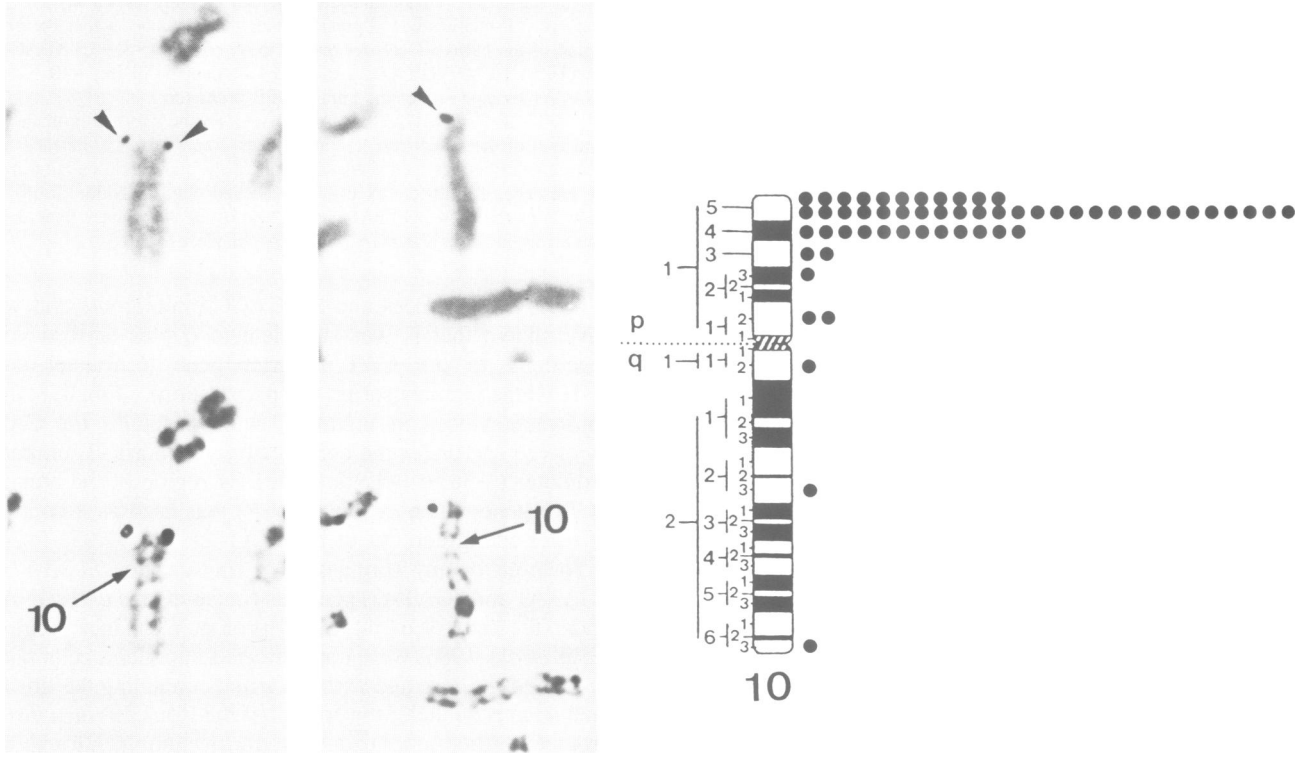

Fig. 3. Localization of hGATA3 to human chromosome 10 by *in situ* hybridization. **left.** Two partial human metaphases showing the specific site of hybridization to chromosome 10. Top: arrow heads indicate silver grains on Giemsa-stained chromosomes after autoradiography. Bottom: chromosomes with silver grains were subsequently identified by R-banding (FPG technique). **Right.** Idiogram of the human G-banded chromosome 10 illustrating the distribution of labelled sites for the hGATA3 probe.

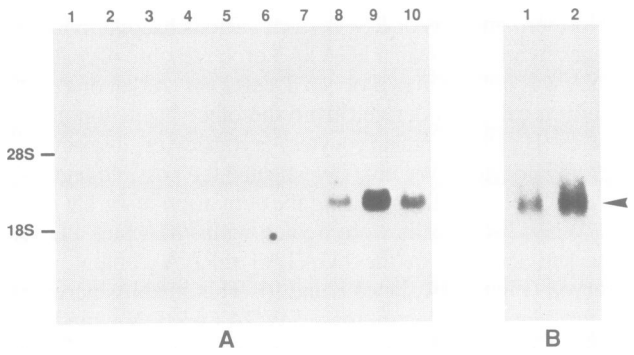

Fig. 4. Expression of hGATA3 mRNA. **A.** Northern blot containing 20 μg of total RNA from the following cell lines was hybridized with a [^{32}P]hGATA3 probe that does not contain the zinc finger domain. Lane 1, K 562 (erythrocytic); lane 2, HEL (megakaryocytic/erythrocytic); lane 3, MEG 01 (megakaryocytic); lane 4, HL60 (promyelocytic); lane 5, U937 (histiocytic); lane 6, an EBV established B-cell line; lane 7, Raji (a B-cell line); lane 8, Jurkat (an α/β T-cell line), lane 9, Jurkat + PMA/PHA; lane 10, Peer (a γ/δ T-cell line). RNA quantity and integrity were checked by staining with ethidium bromide before blotting. **B.** Northern blot analysis of hGATA3 mRNA during T-cell activation. 20 μg of total RNA isolated from resting T-lymphocytes (lane 1) or PMA/PHA-activated T lymphocytes (lane 2) were size fractionated, transferred to a nylon membrane (NYTRAN) and hybridized with a ^{32}P-labelled hGATA3 cDNA probe. The arrow indicates the hGATA3 mRNA.

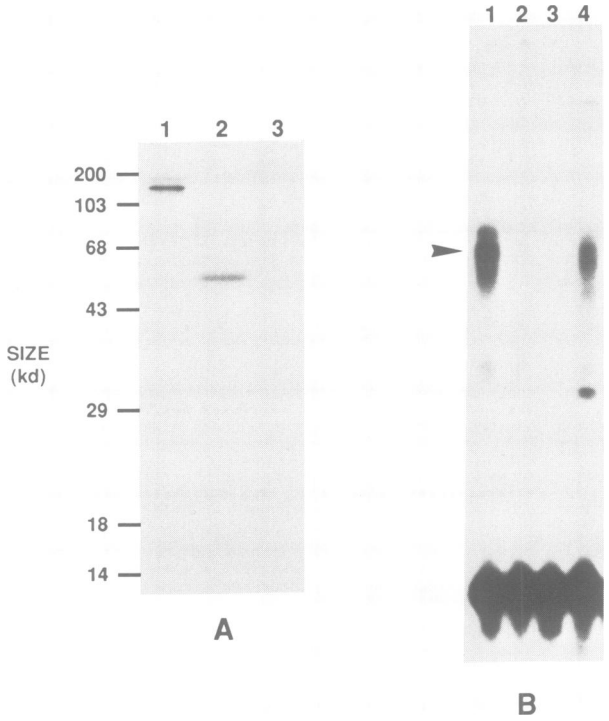

Fig. 5. hGATA3 protein produced *in vitro* specifically binds to the WGATAR motif present in erythroid or T-cell specific regulatory elements. **A.** SDS−PAGE analysis of the *in vitro* translation product of the hGATA3 cDNA clone. 5 μl of the translation reactions without (lane 3) or with hGATA3 mRNA (lane 2) were run on a 12% SDS−polyacrylamide gel and subjected to fluorography. Lane 1 is a control sample [brome mosaic virus RNA (Promega)]. The molecular weight standards used are indicated. **B.** Gel mobility shift assay analysis. ^{32}P end-labelled oligonucleotides that contained either the δE4 region of the human T-cell receptor δ locus enhancer (5′−CTAGAG-TTATCACTTTCTGTTATCAAGTGG−3′) (lane 1) or the −40 hGATA1 binding site of the human GPB promoter (5′−GGGCCTGGAAGATAACAGCTA−3′) (lane 4) were incubated with 2 μl of the translation reaction. Lanes 2 and 3 are ^{32}P-labelled δE4 oligonucleotides incubated with 2 μl of the translation reaction in the presence of a 50-fold excess of unlabelled oligonucleotide [itself (lane 2) or the −40 hGATA1 binding site of GPB promoter (lane 3)]. The arrow indicates the hGATA3-DNA complexes.

the amino acid sequence of hGATA3 (48 kd). This *in vitro* translated product was used in gel mobility shift assays with two different probes. One, from the human glycophorin B (GPB) promoter (Vignal *et al.*, 1990), contains a single AGATAA-binding site, while the other, from the T-cell receptor δ gene enhancer (Bories *et al.*, 1990) contains two TGATAA-binding sites [footprint δE4 (Redondo *et al.*, 1990)]. Each oligonucleotide formed a complex with the [^{35}S]hGATA3-labelled protein (Figure 5B, lanes 1 and 4). The single complex obtained with the δ oligonucleotide, which possesses two WGATAR-binding sites, might be explained by the ability of hGATA3 simultaneously to bind and contact a pair of closely spaced WGATAR motifs as previously shown for hGATA1 and the human $^{A}\gamma$-globin promoter (Martin *et al.*, 1989). The binding to the T-cell receptor δ gene enhancer oligonucleotide could be efficiently competed either by itself or by 50-fold excess of unlabelled GPB oligonucleotide (Figure 5B, lanes 2 and 3).

hGATA1 and hGATA3 can transactivate the glycophorin B promoter, the human TCR δ gene enhancer and a synthetic promoter containing the two WGATAR motifs present in the human TCR δ gene enhancer

As the DNA-binding domain of hGATA1 and hGATA3 was the only homologous region of those proteins, we wished to determine whether those two proteins were able to transactivate T-lymphoid or erythroid regulatory regions which contain WGATAR motifs. A 290 bp DNA fragment, which contained the δE4 region of the TCR δ enhancer and which has been shown to confer T-cell specificity (Bories *et al.*, 1990), was inserted into pBLCAT2 (Luckow and Schütz, 1987) upstream from the *tk−CAT* gene (Figure 6A). We also linked a 95 bp DNA fragment of the human glycophorin B promoter to the CAT reporter gene (Figure 6A). This fragment contained, between −95 and −40, two WGATAR sequences and a CCACC-binding site and we have shown that is was sufficient to confer erythrocytic specificity to the CAT reporter gene (C.Rahuel and P.-H.Roméo, manuscript in preparation). Finally, the construction used in Figure 1, which contains the CAT reporter gene under the control of glycophorin B minimal promoter (−36 to +47) and the δE4 region of the human TCR δ gene enhancer (Figure 6A) was used as a control of *trans*-acting activity of hGATA1 and hGATA3.

To look for any stimulation of transcription *in vivo* by hGATA1 and hGATA3, plasmids containing the two cDNAs under the control of the SV40 promoter/enhancer region (Ellis *et al.*, 1986) were cotransfected into HeLa cells with each reporter gene and with a plasmid containing the firefly luciferase gene under the control of the Rous sarcoma virus promoter. After normalization of transfections, CAT activity showed that hGATA1 and hGATA3 were transactivators of the three constructs (Figure 6B), transactivation levels obtained with hGATA1 being always higher than with hGATA3. Indeed, no transactivation of promoters without

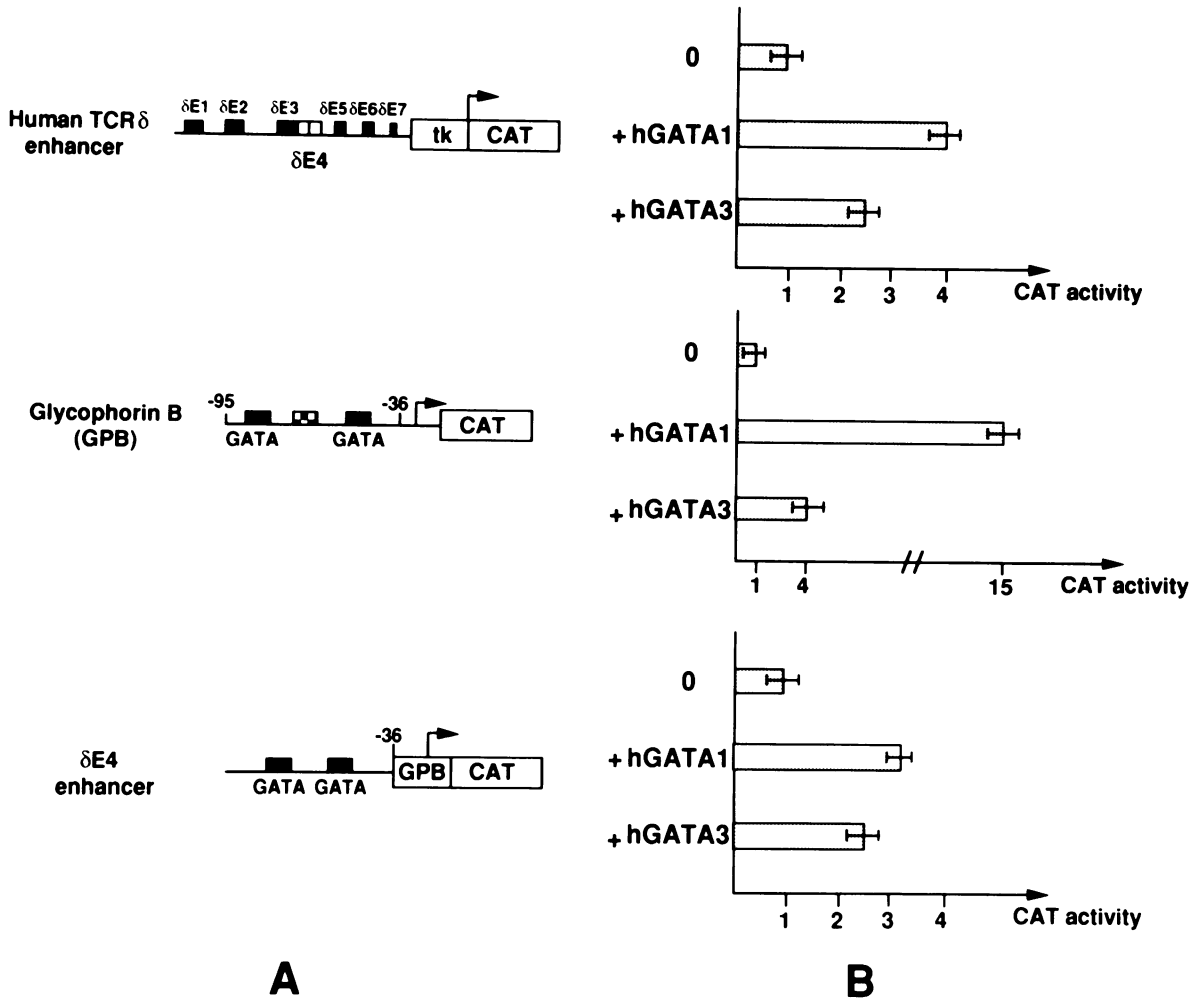

Fig. 6. Transcriptional activation by hGATA1 and hGATA3 proteins. Transactivation analysis was performed by transfecting into HeLa cells,the reporter plasmids shown in Figure 6A wit PECE plasmids expressing either hGATA1 or hGATA3. ▣ indicates a CCACC binding site and ■ indicates GATA binding sites. All the footprints obtained on the human δ enhancer are also indicated. Using half the transfected cells and after normalization of transfections, the transactivation was monitored by CAT activities. The level of CAT expressed above reporter plasmid transfected with PECE alone is shown (Figure 6B) ⊢⊣ indicates the standard errors, the mean being the average of three experiments.

WGATAR sequence (GPB-CAT, pBL CAT 2 or ptk-AGPT-CAT) was observed (data not shown).

As transactivation of the TCRδ enhancer motif with hGATA1 was higher than with hGATA3, we determined the affinity and the concentration of these two factors after transfection of the cloned proteins in HeLa cells. Nuclear extracts from transfected HeLa cells were made and we first ruled out any co-operativity in the binding of hGATA1 and hGATA3 to the two WGATAR motifs present in the δE4 sequence by probe saturation curves (data not shown). We then performed factor titration experiments (Baker *et al.*, 1986). A constant, non-saturating amount of nuclear proteins from HeLa cells transfected with hGATA1 or hGATA3, was incubated in the presence of increasing concentrations of labelled δE4 oligonucleotide. After separation of bound and unbound DNA molecules (Figure 7A), the amount of DNA in each band was determined and represented on a Scatchard plot with the fraction of bound DNA expressed as a function of the amount of retarded DNA. For hGATA1 and hGATA3, the experimental values fitted straight lines (Figure 7B) which were parallel. This indicated that hGATA1 and hGATA3 have a similar apparent equilibrium constant for the δE4 sequence and that the concentration of hGATA1 was 1.5-fold higher than the hGATA3 one. As transfections with higher amounts of SV40 hGATA1 or SV40 hGATA3 gave similar levels of transactivation, this variation might not be related to the observed difference of transactivation.

Discussion

In the study described in this report, we identified a new member of the human GATA family, hGATA3, which is specifically expressed in the T-cell lineage. In contrast to the relatively low degree of sequence homology between the chicken and human GATA1 family members (<45%), the protein sequence of chicken and human GATA3 is highly conserved (>93%). This sequence conservation between species suggested a major transcriptional regulatory function for this factor during T-cell differentiation and/or activation.

Outside of the zinc fingers, hGATA1 bears three internal domains which might interact directly with components of the transcriptional apparatus (Trainor *et al.*, 1990). The hGATA1 amino-terminal domain contains a small acidic

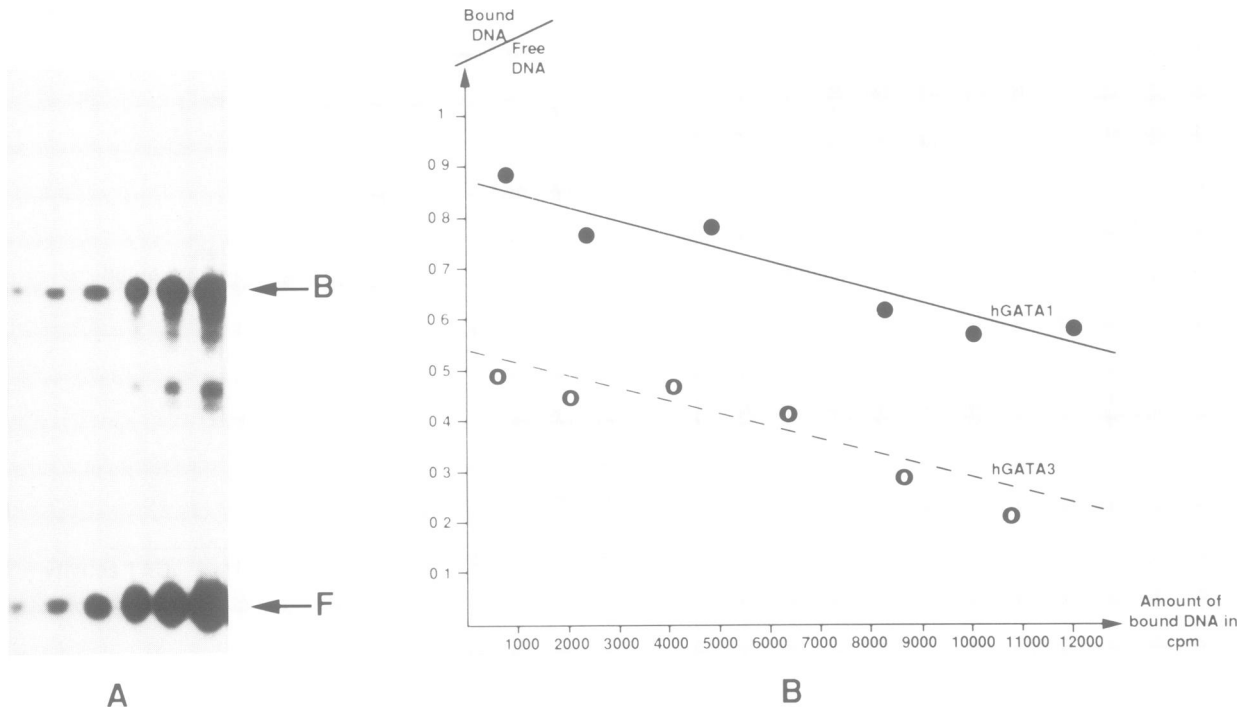

Fig. 7. Expression of hGATA1 and hGATA3 in transfected cells. **A.** Nuclear extracts were made from transfected HeLa cells and used for titration of hGATA1 with the δE4 oligonucleotide. The amount of oligonucleotide varies from 0.016 ng to 0.5 ng. **B** indicates bound hGATA1 and **F** indicates free oligonucleotide. A similar picture was observed with hGATA3. **B.** The ratio **B/F** was plotted against the amount of retarded DNA for hGATA1 (●) or hGATA3 (○) and the two straight lines were drawn by fitting the data using a linear regression.

region characteristic of some transcriptional activators (Ptashne, 1988) and deletion of this region greatly diminished, but did not abolish the transactivating activity of hGATA1 (our unpublished data). As for hGATA3, acidic residues are also found in the amino-terminal region of the protein, but each acidic residue is next to at least one histidine residue. This suggests a possible neutralization of the negative charges of the acidic residues when pH diminished. This hypothesis would imply a low transcriptional activity of hGATA3 in quiescent T-lymphocytes and an increased activity of this factor during the antigen-induced mitosis of T-lymphocytes. This increase, together with an enhancement of hGATA3 gene expression (Figure 4B), would result in a synergy between gene activation and protein activity during T-lymphocyte activation.

The 3′ non-coding sequence of hGATA3 is AT-rich and contains four ATTTA sequences. These features, previously shown to be present in certain lymphokines, cytokines and oncogenes, are supposed to be involved in the instability of these mRNAs (Shaw and Kamen, 1986). Such sequences appeared to be present in the 3′ region of putative T-lymphocyte G_0/G_1 switch genes (Siderovski *et al.*, 1990; Irving *et al.*, 1989) and, as the amount of hGATA3 mRNA also increases during T-cell activation, these sequences might regulate the amount of specific proteins synthesized during the initial stages of T-cell activation.

The haematopoietic expression of hGATA3 mRNA is more restricted than that of hGATA1 mRNA since we could not find any hGATA3 mRNA in non-T-cell haemotopoietic cell lines. Within the T-cell lineage, hGATA3 was found in TCRα/β (Jurkat) and TCR γ/δ (Peer) cell lines, suggesting that hGATA3 expression is not restricted to a specific type of TCR rearrangement. As transcription of the TCR α,β and δ genes is partly regulated by enhancers which contain potential binding sites for hGATA3, we looked for any binding and/or transcriptional activity of hGATA3 to one of these enhancers. hGATA3 was shown to bind δE4, a DNA region of the δ enhancer previously shown to confer T-cell specific gene expression, and hGATA3 was able to transactivate a reporter gene dependent on δE4 for its activity. This finding indicated that hGATA3 could be one of the T-cell specific factors shown to interact with the Tα3 (Ho *et al.*, 1989), Tβ2 (Gottschalk and Leiden, 1990) and δE4, (Redondo *et al.*, 1990) regions of the TCR α,β and δ gene enhancers. These enhancers are active in T-cells but, although the human TCR δ and β enhancers are active in TCR α/β and TCR γ/δ cell lines, the human TCR α enhancer is active only in TCR α/β cell lines. As hGATA3 mRNA was found in both T-cell lines, these results indicate that, at least for the human TCR α enhancer, hGATA3 alone is unable to confer T-cell specific expression. However, the possible binding of hGATA3 to regulatory regions of most TCR genes indicates that, as hGATA1, hGATA3 might co-ordinate the expression of several T-cell specific markers and thus might play an important function during T-cell differentiation and activation.

The existence of two GATA-binding proteins present either in T-lymphocytes or in erythroid cells raised the question of the role of these proteins in the activation of haematopoietic lineage-specific genes. Erythroid and T-lymphoid gene regulatory regions contain, in addition to WGATAR motifs, binding sites for ubiquitous and/or cell-specific factors. The sequence difference of hGATA1 and hGATA3 outside their DNA—binding domain might lead

to a preferential interaction of each protein with an erythroid or T-lymphoid specific combination of other factors. We started this analysis by transfecting, into HeLa cells, either an erythroid-specific promoter or a T-lymphoid specific regulatory region together with hGATA1 or hGATA3. Both proteins were transactivators, but hGATA1 was always more efficient especially when the GPB promoter was used. Our results differ from those reported for the avian clones where similar levels of transactivation were obtained when artificial promoters containing WGATAR motifs upstream from a minimal TATA promoter were co-transfected with cGATA3 or cGATA1 in immortal quail (QT6) fibroblasts (Yamamoto *et al.*, 1990). Furthermore, these synthetic promoters have no enhanced expression relative to the minimal TATA promoter after transfection into erythroid cells (Martin and Orkin, 1990) whereas the GPB promoter is dependent on the WGATAR motifs for expression in erythroid cells (C.Rahuel and P.-H.Roméo, manuscript in preparation). These data suggest that the interactions involved in the transactivation of the GPB promoter are different from those studied with these artificial promoters. As for hGATA3, its poor transactivation on T-lymphoid regulatory regions might be accounted for by the need for another T-lymphoid factor to ensure full activity. The recent finding of an Ets-1 binding site within the human TCR α and β gene enhancers (Ho *et al.*, 1990) strengthens this hypothesis and is currently under investigation.

Materials and methods

cDNA cloning and sequencing

Human bone marrow and T-cell (HUT-78) cDNA libraries in λgt-11 were screened with a PCR-amplified DNA fragment corresponding to the zinc fingers of the human GATA1 sequence under low stringency (40% formamide, 5 × SSC, 1 × Denhardt's, 0.1% SDS at 52°C). The hGATA3 clones obtained were sequenced on both strands by the chain-termination method using T7 DNA polymerase (Sequenase, USB).

In situ hybridization

In situ hybridization experiments were carried out using metaphase spreads from a normal female. Concanavalin A-stimulated lymphocytes were cultured at 37°C for 72 h with 5-bromodeoxyuridine added for the final 6 h of culture (60 μg/ml of medium), to ensure a chromosomal R-banding of good quality. The hGATA3 clone containing an insert of 600 bp in Bluescript (PROMEGA) was tritium labelled by nick-translation to a specific activity of 2×10^7 d.p.m./μg. The radiolabelled probe was hybridized to metaphase spreads at a final concentration of 25 ng/ml of hybridization solution as previously described (Mattei *et al.*, 1985). After coating with nuclear track emulsion (KODAK NTB2), the slides were exposed for 20 days at 4°C, then developed. To avoid any slipping of silver grains during the banding procedure, chromosome spreads were first stained with buffered Giemsa solution and metaphases photographed. R-banding was then performed by the fluorochrome-photolysis-Giemsa (FPG) method and metaphases rephotographed before analysis.

RNA extraction and Northern analysis

Total RNA was extracted from the various cell lines as described (Chomczynski and Sacchi, 1987) and its integrity was checked by electrophoresis on an agarose gel before use. Total RNA (20 μg) denaturation, fractionation by electrophoresis in a formaldehyde−agarose gel, transfer to a nylon membrane (NYTRAN), hybridization with a ^{32}P-labelled hGATA3 cDNA probe and washes were performed as indicated by the manufacturer (Schleicher and Schüell). The filter was then autoradiographed at −80°C for 24 h with intensifying screens.

In vitro transcription and translation of hGATA3 cDNA

hGATA3 cDNA was subcloned into a plasmid vector (pGEM, Promega) that allowed T7 dependent synthesis of RNA. After linearization with *Eco*RI, the RNA was synthesized, checked on a 1% agarose gel and used for *in vitro* translation using rabbit reticulocyte lysate (Promega). Gel mobility shift assay analysis was performed as described (Hope and Struhl, 1985) using the ^{32}P-labelled oligonucleotides described in Figure 5B.

Construction of expression vectors and reporter plasmids

Plasmids which constitutively express hGATA1 or hGATA3 were constructed using the PECE vector (Ellis *et al.*, 1986) where the inserted cDNAs are transcriptionally directed by the SV40 promoter and enhancer. An *Eco*RI DNA fragment containing 120 bp of the 5′ untranslated region, the complete coding region as well as 60 bp of 3′ untranslated region of hGATA1 was inserted at the unique *Eco*RI site of the PECE vector and its orientation relative to the SV40 promoter/enhancer was checked by a *Bam*HI digestion. A *Hin*dIII-*Sst*I DNA fragment that contained part of the coding region as well as 200 bp of 3′ untranslated region of hGATA3 was subcloned into PECE previously cut with the same enzymes. The plasmid obtained was linearized with *Hin*dIII, treated with bacterial alkaline phosphatase and ligated with a *Hin*dIII DNA fragment containing the 5′ untranslated region of the rabbit β globin gene linked to an *Nco*I-*Hin*dIII DNA fragment of hGATA3 containing the complete amino-terminal portion of hGATA3. The resulting plasmid was checked by a *Bam*HI digestion for correct orientation and contained the 5′ untranslated region of the rabbit β globin gene, the complete coding sequence of hGATA3 and part of the 3′ untranslated region of hGATA3 under the control of the SV40 promoter/enhancer.

The GPB reporter gene was constructed by insertion, into a ptkAGPT-CAT vector (Mignotte *et al.*, 1989), of a 95 bp DNA fragment containing the minimal region of the GPB promoter which confers erythroid specificity (C.Rahuel and P.-H.Romeo, manuscript in preparation). The enhancer δ-tk-CAT reporter gene was kindly provided by J.C.Bories (INSERM U.93). The δE4-GPB-CAT reporter gene was constructed using oligonucleotides. It contained the two WGATAR motifs of the TCR δ gene enhancer (footprint δE4) (Redondo *et al.*, 1990) upstream from a minimal GPB promoter (−36 to +47) containing no WGATAR motif.

Transfection of Jurkat and HeLa cells

Transfection into Jurkat cells was performed by electroporation using 10 μg of reporter gene plasmids and 2 μg of an RSV-luciferase plasmid. Co-transfection of reporter genes was performed by electroporation of HeLa cells with the trans-activator, the reporter plasmids together with an RSV-luciferase plasmid. For each transfection analysis, we had to optimize the ratio of transactivator to reporter plasmids to ensure reproducible transactivation. Typically, we used 4 μg of SV40-hGATA1 or SV40-hGATA3 plasmids together with 1 μg of TCR δ enhancer-tk-CAT plasmid, 5 μg of GPB-CAT or 5 μg of δE4-GPB-CAT plasmids. Twenty-four hours after transfection, the cells were harvested and CAT activity was assayed as described (Gorman *et al.*, 1982) using amounts of extract containing identical luciferase activity. Transactivation was quantified by comparison of the CAT activity obtained by co-transfection of the reporter plasmid either with hGATA1 or hGATA3 or with the PECE vector as a mock transactivator.

Micro-nuclear cell extracts were prepared as described (Schreiber *et al.*, 1989) and incubated with the ^{32}P-labelled oligonucleotide corresponding to the WGATAR motifs present in the TCR δE4 gene enhancer. Titration of hGATA1 and hGATA3 with the δE4 oligonucleotide was done as described (Baker *et al.*, 1986) and for quantitative purposes, the retarded and free running DNA bands were cut and counted.

Acknowledgements

We thank Professor J.Rosa for support and encouragement and D.Beaupain, V.Lemarchandel, I.Max-Audit, F.Ajchenbaum, and C.Chabret for constant support. We thank C.Rahuel for providing glycoprotein B promoter, T.Chouard for much help in the *in vitro* assays and J.C.Bories for giving us the δE4 oligonucleotide and the δ-tk-CAT construct. We thank V.Corsetti for help in preparing the manuscript. V.J. and D.B. are supported by the Fondation Pour la Recherche Médicale. This work was supported by grants from the Institut National de la Santé et de la Recherche Médicale, the Centre National de la Recherche Scientifique and the Association pour la Recherche contre le Cancer.

References

Baker,R.E., Gabrielsen,O. and Hall,B.D. (1986) *J. Biol. Chem.*, **261**, 5275−5282.

Bories,J.C., Loiseau,P., d'Auriol,L., Gautier,C., Bensussan,A., Degos,L. and Sigaux,F. (1990) *J. Exp. Med.*, **171**, 75−83.

Chomczynski,P. and Sacchi,N. (1987) *Anal. Biochem.*, **162**, 156–159.
Ellis,L., Clauser,E., Morgan,D.-O., Edery,R., Roth,R.A. and Rutter,W.J. (1986) *Cell*, **45**, 721–732.
Evans,T., Reitman,M. and Felsenfeld,G. (1988) *Proc. Natl. Acad. Sci USA*, **85**, 5976–5980.
Gorman,C.M., Moffat,C.F. and Howard,B.H. (1982) *Mol. Cell. Biol.*, **2**, 1044–1051.
Gottschalk,L.R. and Leiden,J.M. (1990) *Mol. Cell. Biol.*, **10**, 5486–5495.
Ho,I.C., Yang,L.-H., Morle,G. and Leiden,J.M. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 6714–6718.
Ho,I.C., Narayan,K.B., Gottschalk,L.R., Lindsten,T., Thompson,C.B., Papas,T.S. and Leiden,J.M. (1990) *Science*, **250**, 814–818.
Hope,I.A. and Struhl,K. (1985) *Cell*, **43**, 177–188.
Irving,S.G., June,C.H., Zipfel,P.F., Siebenlist,U. and Kelly,K. (1989) *Mol. Cell. Biol.*, **9**, 1034–1040.
Kozak,M. (1984) *Nucleic Acids Res.*, **12**, 857–872.
Kruys,V., Marinx,O., Show,G., Deschamps,S. and Huez,G. (1989) *Science*, **245**, 852–855.
Luckow,B. and Schütz,G. (1987) *Nucleic Acids Res.*, **15**, 5490.
Martin,D.I.K. and Orkin,S.H. (1990) *Genes Dev.*, **4**, 1886–1898.
Martin,D.I.K., Tsai,S. and Orkin,S.H. (1989) *Nature*, **338**, 435–438.
Martin,D.I.K., Zon,L.I., Mutter,G. and Orkin,S.H. (1990) *Nature*, **344**, 444–447.
Mattei,M.G., Philip,N., Passage,E., Moisan,J.P., Mandel,J.L. and Mattei,J.F. (1985) *Human Genet.*, **69**, 268–271.
Mignotte,V., Eleouet,J.F., Raich,N. and Roméo,P.-H. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 6548–6552.
Ptashne,M. (1988) *Nature*, **335**, 683–688.
Redondo,J.M., Hata,S., Brocklehurst,C. and Krangel,M.S. (1990) *Science*, **247**, 1225–1229.
Roméo,P.H., Prandini,H., Joulin,V., Mignotte,V., Prenant,M., Vainchenker,W., Marguerie,G. and Uzan,G. (1990) *Nature*, **344**, 447–449.
Schreiber,E., Matthias,P., Müller-Immerglück,M.M. and Schaffner,W. (1989) *Nucleic Acids Res.*, **17**, 6419.
Shaw,G. and Kamen,R. (1986) *Cell*, **46**, 659–667.
Siderovski,D.-P., Blum,S., Fordsdyke,R.E. and Fordsdyke,D.R. (1990) *DNA Cell Biol.*, **9**, 579–587.
Spencer,D.M.,Hsiang,Y., Goldman,S.P. and Raulet,D.H. (1991) *Proc. Natl. Acad. Sci. USA*, **88**, 800–804.
Takeda,S., Cheng,A., Mauxion,F., Nelson,C.A., Newberry,R.D., Sha,W.C., Sen,R. and Loh,D.-Y. (1990) *Mol. Cell. Biol.*, **10**, 5027–5035.
Trainor,C.D., Evans,T., Felsenfeld,G. and Boguski,M.S. (1990) *Nature*, **243**, 92–94.
Tsai,S.F., Martin,D.I.K., Zon,L.I., D'Andrea,A.D., Wong,G.G. and Orkin,S.H. (1989) *Nature*, **339**, 446–451.
Vignal,A., London,J , Rahuel,C. and Cartron,J.P. (1990) *Gene*, **95**, 289–293.
Wall,L., DeBoer,E. and Grosveld,F. (1988) *Genes Dev.*, **2**, 1089–1100.
Wilson,D.B., Dorfman,D.M. and Orkin,S.H. (1990) *Mol. Cell. Biol.*, **10**, 4854–4862.
Yamamoto,M., Ko,L., Leonard,M., Beug,H., Orkin,S.H. and Engel,D.J. (1990) *Genes Dev.*, **4**, 1650–1662.

Received on January 15, 1991; revised on March 22, 1991